OSU Working Papers in Linguistics 55, 88-116

Lexical Effects in the Perception of Obstruent Ordering[1]

Matthew J. Makashay
makashay@ling.ohio-state.edu

1. INTRODUCTION

Cross-linguistic study of obstruent metathesis (Bailey 1970; Ultan 1971; Silva 1973; Hock 1985; Hume 1998, 2001; Steriade 2001) has attempted to understand this process. Many accounts delve into human auditory perception to aid in the explanation of these seemingly complex patterns of re-ordering.[2] Hock proposed a perceptual motivation for the preference of the ordering fricative-stop word initially (prevocalically), as there are clearer transitions between stops and vowels than between stops and fricatives.

Hume provided a systematic perceptual account for both place and continuant metathesis. The strengths and weaknesses of cues in different consonantal positions could give impetus for a consonant to switch to a position with better cues if there is little detriment to the other consonant, resulting in overall better perception of the sequence. An example of this process is the metathesis of /VkpV/ to [VpkV] in Kui (Hume 2001), shown in Table 1. The explanation for this place re-ordering involves the strong burst for

[1] Thanks to Elizabeth Hume, Keith Johnson, Mary Beckman, Michael Broe, Lisa Shoaf, Amanda Miller-Ockhuizen, Thomas Stewart, Kiyoko Yoneyama, Jeff Mielke, Steve Winters, Misun Seo, Huang Tsan, Georgios Tserdanelis, and Peggy Wong for their input. This material is based upon work supported under a National Science Foundation Graduate Fellowship.

[2] Articulatory explanations, which will not be focussed on in this paper, have also been proposed. For example, Bailey's observation that metathesis can result in apicals following nonapicals (and nondorsals following dorsals) led to his proposal that metathesis may be driven by a preference for a "natural" physiological ordering.

prevocalic velars (vs. the weak burst for labials) and the good vowel formant transitions to labials in postvocalic position (vs. good formant transitions to velars after certain vowels). The /k/ in postvocalic position ([VkpV]) is not as good perceptually as it would be in prevocalic position (Winters 2001). The gain of cues for /k/ by switching to prevocalic position with its burst more than compensates for the loss of linearity, since the consonant might have been perceptually lost otherwise. /p/ is relatively robust in its new postvocalic position, barely losing any cues besides its weak burst. Thus, a cluster containing /p/ and /k/ has the best overall perceptual cues identifying each segment if they are placed in this "optimal" order as [VpkV]. Conversely, the cluster [VkpV], with poorer cues overall in identifying each segment, is "non-optimal."

Table 1. Example of [kp] metathesis in Kui (Hume 2001)

Verb Stem	*Past* -te	*Pres. Part.* -pi	*Gloss*
ah-	ahte	ahpi	'to hold'
lek-	lekte	lepki	'to break'

The metathesis of stops and fricatives is accounted for by the fact that fricatives have strong internal cues (as well as external), while stops only have external cues such as vowel transitions or bursts. Therefore, a stop will gain better perceptibility if it can switch to a position with better external cues. Steriade (2001) states that sibilant-stop (ST) and stop-sibilant (TS) sequences are confusable, citing evidence from Pickett (1958) and Fay (1966) which she claims indicates that the linear order of adjacent consonants that share manner features (such as obstruency and continuency) is highly non-salient. Incorporating this with Hume's perceptual account forms her hypothesis that metathesis could arrive from listener error: Mishearing an obstruent cluster of a stop and a sibilant would be constrained by perceptual optimization, resulting in a sibilant-stop cluster if prevocalic, otherwise in a stop-sibilant cluster (if postvocalic). A stop will change positions to gain a good burst prevocalically. If there is no prevocalic position, it will switch to a postvocalic position to gain vowel formant transition cues.

However, while they are certainly less systematic than the optimal result, both prevocalic TS[3] and postvocalic ST[4] metathesis results have occurred historically across languages, but very rarely. Silva (1973) proposes that the cases involving metathesis to stop-sibilant only occur in languages with affricates[5], so speakers are previously accustomed to stop-sibilant sequences. Even though some data go against perceptual optimization, since there are so few examples, they could be the result of chance misperceptions that were learned. Although one cluster ordering is more "optimal" than

[3] Nakao (1986) has these examples from OE: ā[sk]ian > ā[ks]ian 'ask' (but also ā[ks]ian > ā[sk]ian), and a[sk]e > a[ks]e 'ashes'. Silva (citing Collinder 1960) notes a case in the development of Lappish into Mordvin: boaške 'the small of the leg' > pukšo 'the thick flesh; thigh, buttock'.

[4] Nakao has more examples from OE: tu[ks] > tū[sk] 'grinder' and wæps > wæsp 'wasp'. Silva has a case in the change from OE to ME: do[ks] > do[sk] 'dusk'.

[5] Which is true for all the languages in footnote 3.

another by having better perceptual cues, it just has a better chance of becoming an output, not the only chance.

Perceptual cues may not be the only factor involved in metathesis, as lexical effects may influence the ordering of obstruents as well. Much of the evidence used to support cue saliency of different consonants in various positions comes from perceptual studies. However, many of these studies do not take into account lexical effects, such as word frequency and phonotactics. These effects have been shown to affect perception (Luce 1986, Luce and Pisoni 1998, Pitt and McQueen 1998, Vitevitch and Luce 1999, Frisch et al 2000). If metathesis is motivated by the re-ordering of consonants to provide overall better cues to their identity, this should influence the cluster inventory of the lexicon. However, the lexicon may influence the perception of consonant clusters toward orders that occur more frequently.

This study will attempt to determine whether perceptual cues, lexical effects, or both influence the ordering of obstruents in American English. After a brief overview of possible acoustic cues that perceptually distinguish one consonant from another in English medial obstruent clusters, previous perceptual experiments on these clusters will be summarized, then the frequency of medial obstruent clusters that occur in English words will be examined, followed by a report of a perceptual experiment factoring in both acoustic and lexical information. English was chosen as the language of study because of the extensive previous research on it both phonetically and lexically. While phonetic research has been conducted cross-linguistically, only a handful of languages have enough analyzed corpora to yield spoken word frequencies, word familiarities, and other aspects of the lexicon that may influence the recognition of a word. In order to determine if different acoustics have a perceptual effect, any lexical factors need to be taken into account.

2. ACOUSTIC CUES

Before hypothesizing about how perception plays a role in metathesis, the assumed acoustic basis of the perceptual cues needs to be discussed. For the obstruents in question, stops and fricatives, the acoustic cues vary by phone and also by position within the cluster. Due to restrictions of the English lexicon—some clusters do not occur in enough words to have a large enough sample for testing—only stop-stop clusters composed of /p/, /t/, and /k/, and fricative-stop and stop-fricative clusters (fricative%stop clusters) composed of /p/, /t/, /k/, and /s/ will be investigated in the study.

2.1. Stop-stop clusters

Much research has been performed on determining the acoustic cues for stops (Delattre et al 1955, Öhman 1965, Blumstein and Stevens 1979, Kewley-Port 1983, Lahiri et al 1984, Stevens 1989, Wright 2001, for example). One conclusion that may be drawn for American English is that labials and velars have better vowel formant transition cues in postvocalic position than coronals. This is due to acoustic indications for place, such as the lowering of formants for labials, and the "velar pinch" for velars after front vowels, which are illustrated in Figure 1. While American English /t/ may have a good cue to its identity by being glottalized, (sometimes also surfacing as a glottal stop), some talkers glottalize their vowels, rendering this cue useless in many cases.

Another conclusion that may be drawn from this research is that velars and coronals have better burst cues in prevocalic position than labials. Coronals have a higher frequency of burst energy than velars and labials, and velars may have two or more bursts, shown in Figure 1. But labial bursts have weaker, more diffuse energy than velars or coronals. Vowel formant transition cues are also good in prevocalic position, being almost mirror images of postvocalic transitions, but can be obscured by aspiration after voiceless stops in English. Another factor is that labials have shorter transitions to the following vowel, which makes them more likely to be masked than velars or coronals.

Overall, prevocalic position is better than postvocalic position for the identification of stop place (Blumstein and Stevens 1979, Wright 2001). One indication of this is that CV syllable structures are preferred over VC cross-linguistically.

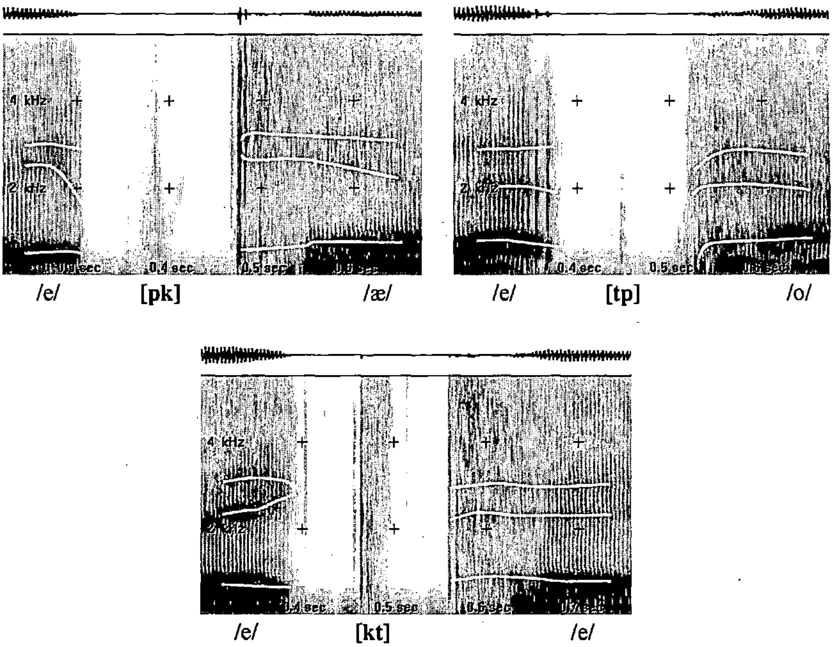


Figure 1. Note the falling of (traced) formant frequencies for postvocalic [p], the pinch of formants F2 and F3 for postvocalic [k] (for front vowels), and the glottalization that marks a following [t]. For the prevocalics, notice the strong double burst for [k] and the burst for [t], while [p]'s burst is very weak. Although the vowel formant transitions are obscured by aspiration, they are roughly symmetrical to those preceding the corresponding stop in postvocalic position.

Prevocalic position has burst information as well as transition cues, while postvocalic stops may be unreleased. Also, American English /t/ is not glottalized prevocalically, so it has good identification cues from its burst and following formant transitions. Figure 1 displays token utterances of an American English female illustrating the acoustic cues mentioned above for postvocalic and prevocalic [p], [t] and [k].

Based on the acoustic evidence discussed, predictions can be made on optimally positioned stops, and the clusters they compose. An optimal postvocalic stop will be indicated by a preceding '>' that symbolizes perceptually good, right-pointing cues (e.g. >[k]), and an optimal prevocalic stop will be indicated by a following '<' that symbolizes perceptually good, left-pointing cues (e.g. [k]<). As illustrated in Figure 2, postvocalically, labials and velar stops have better place cues than coronals, so >[p] and >[k] outrank [t]. Prevocalically, coronals and velars have better cues than labials, so [t]< and [k]< outrank [p]. Furthering the symbolism, optimal clusters will be encased in '> <'. For example, since >[p] has good postvocalic cues and [t]< has good prevocalic cues, the cluster they compose will be represented as >[pt]<.

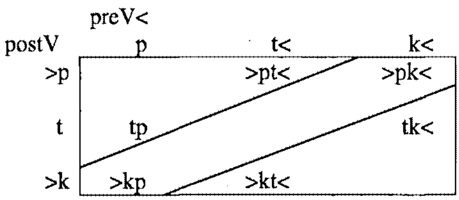


Figure 2. Predicted American English cluster perceptual goodness outcomes based on rankings of acoustic cues. Prevocalic position has more cues than postvocalic position. Compare the predicted cues within each diagonal section.

A comparison of the mirror-image pairs within each diagonal section of Figure 2 can be used to determine which cluster is better perceptually, based on the acoustics. A list of each optimal/non-optimal pair is shown in Table 2. A cluster with a greater number of good post and prevocalic cues would be perceptually stronger (optimal) than a cluster with fewer cues (non-optimal). For example, >[pk]< is optimal when contrasted with >[kp] due to the number of better cues. Non-optimal clusters have been stripped of any cue symbols for easier readability.

Table 2. Optimal and non-optimal stop-stop clusters in intervocalic position

optimal	non-optimal
>pt<	tp
>pk<	kp
>kt<	tk

2.2. Fricative-stop vs. stop-fricative intervocalic clusters

Fricatives always have internal fricative noise frequencies as a place cue, whether they are pre or postvocalic. Stops, on the other hand are better in prevocalic position, as shown above. Determining the ordering of intervocalic fricatives and stops is better when the stop follows the fricative because there is separation between frication and burst

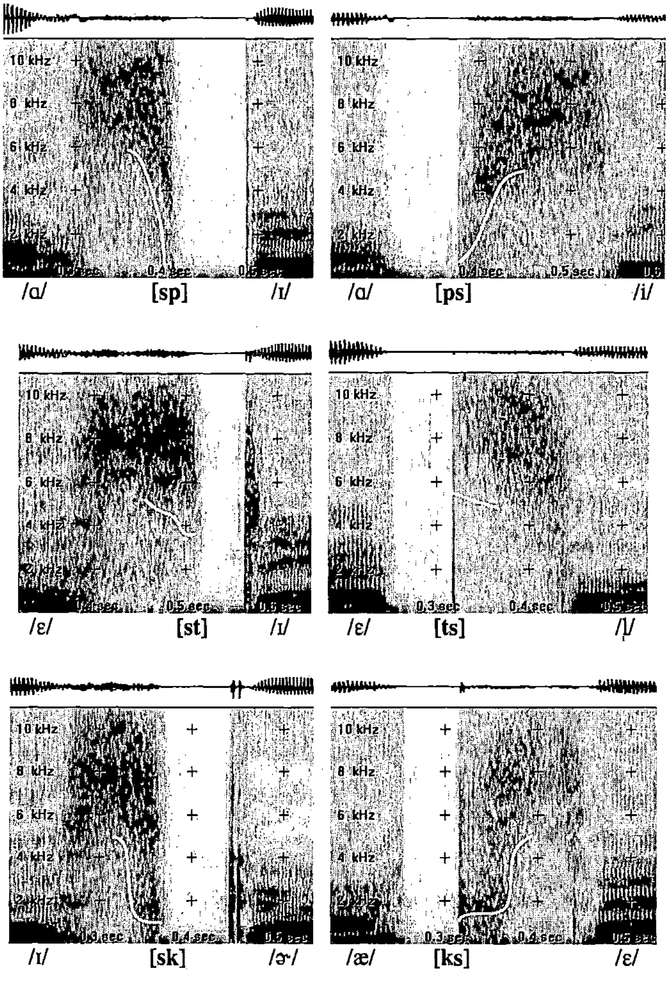


Figure 3. [s] always has internal noise as evidence to its place, but also has transitions to and from stops of a different place. White lines underscore the lowest peak of spectral energy in [s].

(frication followed by silence, then the stop burst) as opposed to when the stop precedes the fricative (silence followed by the burst, then frication) (Wright 2001). Although forward masking can occur for up to 75 to 100ms after the fricative ends, which could still overtake many following stop bursts, the burst of a preceding stop is well within the 50ms limit for backward masking (Yost, 1994). There are also stop place cues in the transitions to or from a neighboring fricative. Figure 3 displays examples of clusters with [s] and [p], [t], or [k]. A listing of the optimal clusters with prevocalic stops after the fricative, and non-optimal clusters with postvocalic stops before the fricative are shown in Table 3.

Table 3. Optimal and non-optimal intervocalic clusters with fricatives and stops

optimal	non-optimal
>sp<	ps
>st<	ts
>sk<	ks

3. PERCEPTION

3.1. Obstruent place of articulation

Testing the perceptual goodness of the acoustic cues of obstruents has not been a straightforward affair. Presenting consonants in words and non-words to subjects at low volumes and with background noise to examine the listening errors has resulted in different rankings of place salience. Some of the differences in the findings can be explained by the inventory (whether there were fricatives or voiced stops), whether bursts were present or not, the vowels used, and the fact that different talkers have different oral cavities and articulation patterns. The following are a sampling of salience rankings from studies of place perception in English:

(1)	Miller and Nicely (1955):	CV coronal > dorsal, labial	
(2)	Wang and Bilger (1973):	CV labial, coronal > dorsal	
		VC coronal > labial > dorsal	(burst not specified)
(3)	Hume et al (1999):	CV dorsal, labial > coronal	(English and Korean)
(4)	Winters (2001):	CV labial, dorsal > coronal	
		VC labial > coronal > dorsal	(burstless)
(5)	Wright (2001):	CV labial > coronal, dorsal	(burstless)
		VC labial > coronal, dorsal	(burstless)

As there is much variation among these and other perceptual consonant ranking studies, the apparent overall trend of the ranking positions, modified by American English acoustics and the particular talker's speech characteristics, led to a set of optimal perceptual clusters identical to the optimal acoustic clusters in Table 2.

3.2. Segment ordering

One perceptual perspective that accounts for temporal disordering of stops and fricatives, as Steriade (2001) claims, is auditory streaming (Bregman 1990). In this account, the high frequency of fricative noise is perceptually far enough away from vowel formants to separate speech into separate streams—one containing fricatives, and one containing vowels and other sonorants. Temporal ordering across streams is difficult, as there are few acoustic cues that line up in both streams: Vowel formant transitions that give stop place cues are lower than frication frequencies, and a stop before a fricative could have its burst masked by the fricative. Switching the stop to a position in which it has a strong burst would bring it into the fricative stream, which is the expected result of metathesis prevocalically ([STV]). A stop that is preobstruent or phrase final may be unreleased, resulting in lack of evidence for it following the fricative. This may increase its chance of being ordered before the fricative ([VTS{C,#}]), instead of after ([VST˺{C,#}]), similar to the patterns observed in Faroese and Lithuanian (Seo and Hume 2001).

There also have been perceptual studies that have observed metathesis errors by subjects listening to clusters, and a few have tested aspects of the linear ordering of segments in clusters. One example is Pickett (1958), as noted above, which tested the perception of consonant clusters in noise (using flat noise with a signal-to-noise ratio at -4 dB and at +6 dB, and low-frequency noise with a spectrum slope of -12 dB per octave with a signal-to-noise ratio at -30 dB). Only the final consonant cluster syllables—bVCC—had alternate sibilant-stop pairs ([ts] and [st], as well as [ks] but not its pair). The largest reported listener error for the coronal pairs in the -4 dB flat noise was [ks], and the second was perceptual metathesis[6] (with a higher rate for [st]). In the low-frequency noise, the largest error for [ts] was [ks], followed by [ls] and then [st]. [st] did not have as high error rates. However, some of the stimuli used were actual English words, while others were not. English words containing [i], [ɑ], and [o] formed by bVks (*beaks* and *box*) have spoken and written frequencies over four times higher than bVst words (*beast*, *bossed*, and *boast*), which had over 16 times higher written frequencies (but similar spoken) than bVts words (*beets/beats*, and *boats*). Also, the responses were forced choice, preventing alternatives such as [sk] or [p].

Fay (1966) investigated subjects' temporal resolution of voiced non-plosive pairs (including nasals, fricatives, and liquids) and pure-tone pairs in noise, with no surrounding context. Staggered onsets with different lag and lead times of voiced non-plosives were played to subjects' right ears with equal offset times. The task was to determine which consonant came first, with onset lead and lag times of 70, 50, 30, 10, and 0 ms. Although stops were not used, nasal-fricative sequences seemed to break the expected pattern of fricative-stop clusters being easier to perceive word initially, with the timing of nasal-fricatives ([nz] and [nð]) perceived correctly more often than corresponding fricative-nasals. (Though there is a bias in hearing [ð] first in the [nð] pair.) The median scores were 100% for seven phoneme pairs out of twelve, and three pure-tone pairs out of four, but half the phoneme pairs had better temporal resolution than

---

[6] The term "perceptual metathesis" is used to indicate that the process is not incorporated in the grammar, as it was heard in manipulated laboratory speech and is not used systematically.

pure-tone pairs. Fay explains this by suggesting that linguistic experience gave subjects higher accuracy on temporal resolution of individual phoneme pairs than less natural pure-tone pairs.

Bond (1971) performed a perceptual experiment on the perception of stop-sibilant and sibilant-stop clusters ({p,t,k}s, s{p,t,k}) inter-vocalically and postvocalically in English words. White noise was added to spoken words to attain different signal-to-noise ratios of 0 dB, +12 dB, and -6 dB. Subjects were told before the test that some of the words were unusual, and were shown them. The test was presented twice, with responses written the first time, and spoken the second. Bond found that the most common error is perceptual metathesis, with the sibilant-stop clusters perceived correctly less often than the corresponding stop-sibilant. However, in inter-vocalic position the sibilant-stops were heard as stop-sibilants less often than the stop-sibilants were heard as sibilant-stops. But the sibilant-stop tokens used (*Caspian*, *blister*, and *asking*) have higher Kucera-Francis written frequencies and Brown verbal frequencies than the stop-sibilant tokens (*Capsian*, *blitzer*, and *axing*), so the result could be a lexical frequency effect, as the subjects would expect to hear the more common words.

Although many studies have tested the perceptual cues of obstruents, it appears that some of the results are contradictory. Place salience rankings are not in agreement with each other, and neither are the preferences for stop-fricative orderings. One possible explanation for this variation could be lexical effects, as the number of words that contain an obstruent sequence may be as important for recognition as its perceptual cues.

4. LEXICAL COUNTS AND FREQUENCIES

4.1. Counts

Given the acoustic and perceptual ranking of obstruent clusters in Table 2 and Table 3, the prediction is that the number of optimal clusters is greater than the number of non-optimal clusters in English words. All else being equal, if there are diachronic changes due to misperceptions, the optimal clusters should be more stable and therefore be in more lexical entries than the non-optimal clusters. These predictions are represented in Table 4.

Table 4. Predictions of lexical counts of words with optimal and non-optimal intervocalic clusters.

Number of words with optimal cluster	>	Number of words with non-optimal cluster
>pt<	>	tp
>pk<	>	kp
>kt<	>	tk
>sp<	>	ps
>st<	>	ts
>sk<	>	ks

Tallying the number of English words with medial obstruent clusters listed in the CELEX lexical database (Baayen, Piepenbrock, and Gulikers, 1995) results in totals that mostly support these predictions. For both all English words (Figure 4) and only monomorphemic English words (Figure 5) the optimal clusters >/pt/<, >/kt/<, >/sp/<, and >/st/< occur more often than their non-optimal counterparts, while >/pk/< does not appear to be much more common in words than /kp/. /ks/, the only cluster in English with one alphabetic letter, *x*, appears in more words than optimal >/sk/<, and also occurs in more words than all other non-optimal clusters combined. Because of orthography, and large lexical representation, /ks/ may be a better perceptual unit for English speakers than >/sk/<.

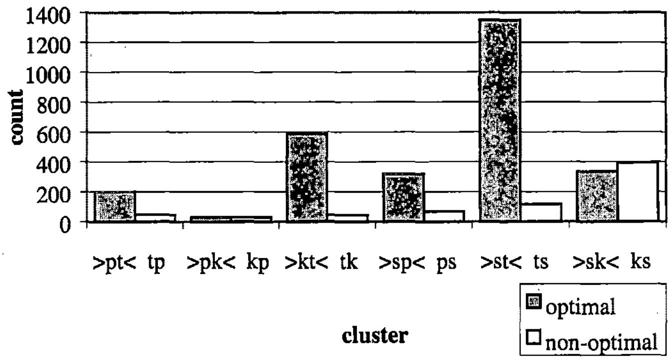


Figure 4. Word count of *VCCV* English words in the 52.5 thousand word pronunciation dictionary in CELEX.

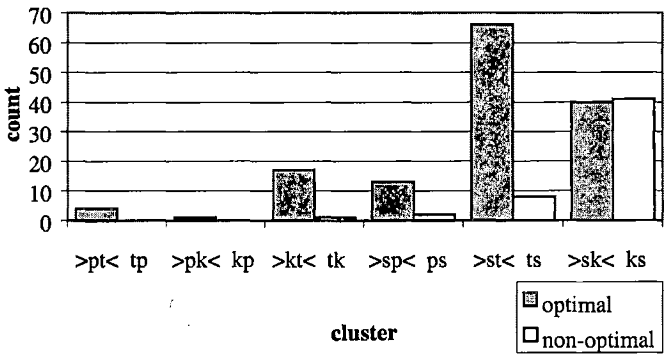


Figure 5. Word count of monomorphemic *VCCV* English words in the 52.5 thousand word pronunciation dictionary in CELEX.

4.2. Frequencies

While the ratio of words with optimal and non-optimal clusters in the lexicon supports the prediction, the usage of these words could affect their perceptibility. Clusters that are spoken more often will be heard more often, perhaps tuning the perceptual system towards detecting their cues more accurately than for clusters heard less often (Frisch et al 2000). Perception of words is affected by their frequency of occurrence, the number of neighboring words that are phonetically similar to them, the predictability of the segment sequences, and how familiar they are to the listener, among other factors (Pollack et al 1959, Savin 1963, Luce 1986, Luce and Pisoni 1998, Pitt and McQueen 1998, Vitevitch and Luce 1999, Frisch et al 2000). Pragmatic, semantic, and syntactic information also play a role. So, although optimal clusters occur more often in the lexicon, are they also spoken and heard more often than non-optimal clusters?

Summing the frequencies of occurrence of words with the clusters of interest in the COBUILD word corpora in CELEX yields the results shown in Figure 6 for spoken and written frequencies. Values for spoken frequencies alone are proportional to the overall sum of spoken and written. Patterns of frequencies of occurrence, and counts of word medial obstruent clusters are highly similar to each other, e.g. /ks/ occurs more than >/sk/<, and >/pk/< is barely more frequent than /kp/. This predicts that optimal clusters, which have better cues, will also be heard more frequently than non-optimal clusters, aiding in their perception. However, /ks/ may compete with >/sk/< for which is better perceptually—/ks/ occurs more often, but >/sk/< has better cues.

Overall, lexical counts and frequencies are consistent with the perceptual account of optimality. But, since there are cases in which they are at odds with each other, both perceptual and lexical effects were controlled for in the experiment.

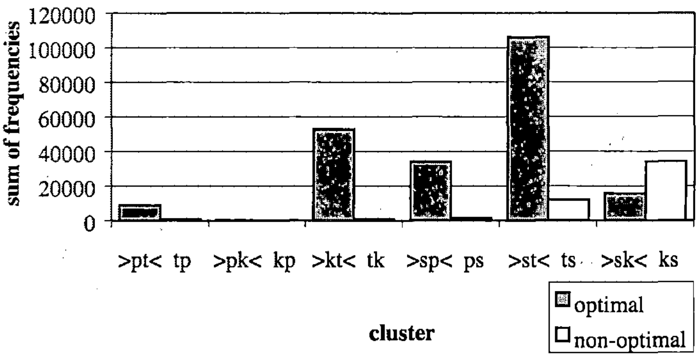


Figure 6. Sum of spoken and written frequencies of *VCCV* English words in the COBUILD 16.6 million written word corpus and in the COBUILD 1.3 million spoken word corpus in CELEX.

5. EXPERIMENT

To test whether the perceptibility of clusters depends on acoustic and lexical information, a lexical decision and repetition task based on natural, spoken American English words containing word-medial obstruent clusters ((C)*VCCV(C)*, e.g. *napkin*) was performed. Actual spoken words were used instead of nonsense syllables, since in ordinary speech communication listeners are trying to perceive meaning in words transmitted acoustically, not the order of consonants. The number of lexical items that contain a cluster may influence the perception of that cluster, which would be the result of word frequency and neighborhood effects (Luce 1986, Luce and Pisoni 1998, Pitt and McQueen 1998, Vitevitch and Luce 1999, Frisch et al 2000). Thus, words with high and low spoken and written frequency were used. Non-words that are the metathesized pairs of these words (e.g. [nækpɪn] for *napkin*) tested if subjects perceived the obstruent order correctly or not. If they heard the order correctly, the subjects would have decided that the token is a non-word ("*nakpin*"). If the subjects did not perceive the obstruents correctly and perceptually metathesized them, they would have decided that the token is a word (*napkin*). This is similar to a mispronunciation detection task.

The types of obstruent clusters used in the experiment had place differences for the stop-stop clusters ({p,t,k}-{p,t,k}), and continuant (and sometimes place) differences for the fricative-stop clusters (s-{p,t,k}) and stop-fricative clusters ({p,t,k}-s). The stop clusters tested the effects of place on confusability, while the clusters with fricatives and stops tested if the perceptual optimization hypothesis is supported by confusions resulting in metathesis only surfacing as fricative-stop before a vowel. Steriade's claim that the sharing of manner features corresponds with the non-saliency of the linear order of two consonants was also tested, and would be supported if the clusters with fricatives and stops metathesized less than clusters with only stops when the subjects identified the stimuli as words or not.

5.1. Predictions

A non-word token will more likely be perceptually metathesized to form a real English word if the resulting cluster is optimal (controlling for word frequency and neighborhood density), as demonstrated in Figure 7. If a subject hears a non-word token with a non-optimal cluster, there may be confusion as to the ordering of the cluster. If there is a real word that can be formed by metathesizing to an optimal cluster, the subject may decide the real word was what was actually spoken. Or there may be confusion, causing a longer reaction time, but the subject finally decides that the token is not a word. If a subject hears a non-word token with an optimal cluster, there should be little confusion as to the ordering of the cluster. The subject should quickly decide that the token is not a word. Subjects' responses should parallel the outcome from historical and grammatical metathesis: non-optimal clusters switching to optimal, and optimal clusters being maintained should be the overwhelming pattern.

Perceptual improvement effect	
non-optimal => optimal	optimal ≠> non-optimal
False alarm or slow	Correct rejection & fast
Stimulus => Percept ks => >sk<	Stimulus ≠> Percept >sk< ≠> ks
[wɪksi] => whiskey	[tæski] ≠> taxi

Figure 7. Non-words containing clusters with "poor" acoustic cues like [wɪksi] should metathesize to English words with "good" acoustic cues. Non-words containing clusters with "good" acoustic cues like [tæski] should not metathesize to English words with "poor" acoustic cues.

The expected word frequency effect will be controlled for by balancing the overall group frequencies between optimal and non-optimal word pairs, but this cannot be done for the cluster frequencies themselves. As discussed in section 4, the number of words that contain a particular cluster and how frequently these clusters are used in speech can vary widely between optimal and non-optimal pairs, in some instances by a factor of ten. The result may be that there is a cluster frequency effect, which would be demonstrated by the listeners' faster reactions to or higher accuracy for words that contain high frequency clusters than for words with low frequency clusters. If this were the case, the results should be the same as for the optimality condition, since for most cluster pairs the optimal one also is the most frequent. The two cluster pairs that would go against this pattern are >[pk]< and [kp], which have roughly the same count and frequency and therefore should show no effect, and >[sk]< and [ks], in which the non-optimal cluster has a higher count and frequency and therefore would aid better performance.

5.2. Methods

Stimuli The stimuli were composed of targets and two types of foils. In order to minimize possible word-level stress effects, the attempt was made to only use words with the same stress pattern. The CELEX database was used to find trochaic[7] English words that also met the required cluster criteria. The targets were non-words produced by metathesizing the medial obstruents in these words. For example, [tæski] was created by metathesizing the [ks] in *taxi*, and [rɛtpaɪl] was created by metathesizing the [pt] in *reptile*. Other English words with medial obstruents were used for real word foils (e.g. *ritzy* and *dropkick*). The non-word foils had zero phonological neighbors (by addition, subtraction, or substitution of a phone) and came from the substitution of medial obstruent clusters into English words with zero frequency and zero neighbors. For example, [flæspʌn] was created by substituting [sp] into *flashgun* and [haʊtkɔg] was created by substituting [tk] into *housedog*.

[7] As there were not enough trochaic words to provide an adequate number of tokens in each cluster group, some compounds with primary stress on the first syllable were used as well.

The list of 120 targets and their lexical sources appears in Appendix A, grouped by the resulting metathesized clusters. The clusters used are >[pt]< and [tp], >[pk]< and [kp], >[kt]< and [tk], >[sp]< and [ps], >[st]< and [ts], and >[sk]< and [ks]. Ten words per cluster for twelve clusters yield 120 targets. Cluster pairs (e.g. >[sk]< and [ks]) are balanced so they have similar word onsets and offsets and have similar total frequencies of occurrence.

The 120 English word foils are shown in Appendix B. These words have the same clusters used in the metathesized tokens, and have similar word onsets and offsets. However, due to the limited number of English words with the VCCV pattern, some of these tokens have consonants adjacent to the medial obstruents. Nonetheless, the word foils were constructed to be phonetically similar to the targets.

The list of 120 non-word foils and their lexical sources are shown in Appendix C. None of the English source words have a medial obstruent cluster used in the experiment. These words have a zero frequency of occurrence, and have a neighborhood density of zero (since no word in the CELEX database was phonemically similar based on the additions, subtractions, or substitutions of a single segment). Similarly, the non-words created by substituting the medial cluster with one of the 12 clusters used in the experiment have zero neighbors as well. Each obstruent in the substituting cluster was the result of a change in place or manner (and optionally voicing)[8] of the original obstruent. For example, the medial cluster in *squad car* was changed to [st] to make the non-word foil [skwɑs.tɑr] by changing the manner (and voicing) of /d/ to yield [s], and the place of /k/ to yield [t]. These non-words vary in degrees of "word-likeness" as an attempt to increase the difficulty of separating words from non-words. The first two tokens for each cluster do not contain a word for either syllable (e.g. [slut]-[pælv] from *sluice-valve* in the [tp] group). The following three tokens contain a word only in the second syllable (e.g. [stɑt]-*par* from *stockcar*). The next three tokens contain a word only in the first syllable (e.g. *pit*-[pot] from *pigboat*). The final two tokens contain words in both syllables (e.g. *greet-pun* from *grease-gun*).

In total, there are 360 stimuli in the lexical decision task: one-third are real word foils the subjects should reply YES to, one-third are non-word foils the subjects should reply NO to, and the remaining third are metathesized targets the subjects may reply YES or NO to depending on whether or not the tokens are perceptually metathesized to form real words. The foils also helped determine if the subject performed the task correctly.

Talker The talker was a female native Ohio English speaker with phonetic knowledge and no known speech or hearing disorders.

Procedure for talker Randomized lists of the tokens were read at a steady rate until three accurate repetitions were achieved. For the non-word targets and foils, the intended pronunciation was elicited by displaying the English word, followed by the cluster to substitute word medially:

[8] Some of these non-words resulted from a change only in voicing, as there were not enough English words that satisfied the 0 frequency/0 neighbors constraint.

(6) pizza
st

root beer
pk

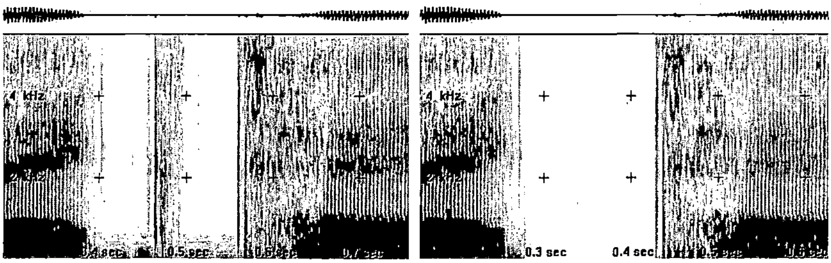
The talker would say the English word, followed by the non-word with the substitution. Stimuli were recorded onto DAT-tape using a Shure SM10A head-mounted microphone, and re-digitized to create computer soundfiles at 22.05 kHz.

The digitized words were edited to ensure that the soundfiles began with word onset and ended with word offset. Since the talker did not release all postvocalic stops in stop-stop clusters (and since this occurs in natural speech), the amplitudes of all stop-stop closures were reduced to zero, even if there was no detectable burst. This is demonstrated in Figure 8. Clusters with stops and fricatives were unaltered.

Two phonetically trained researchers naïve to the purpose of the experiment judged the accuracy of the pronunciations based on a provided list of transcriptions. Items that did not score 4 or above on a 5-point goodness scale by both judges were discarded, which only occurred for less than 1% of the cases.

Listeners The listeners were 30 native Ohio English speakers who were undergraduates at The Ohio State University with no known speech or hearing disorders. 20 of them heard the stimuli at a comfortable listening level (CLL group), and 10 of them heard the stimuli nearly at their speech reception threshold (SRT group). The listeners received partial course credit for their participation.

Procedure for listeners The experiment involved two tasks—an auditory lexical decision task (Goldinger 1996) and a repetition task. The purpose of the repetition task was to confirm that if a subject decided a metathesized target was a word, then the subject had indeed metathesized it to the intended word, and did not make a different error to create some unrelated word. For example, if the subject heard the target [mɪskɚ], decided it was a word, and stated it was *mister*, then it would not be treated as a



$[kt^h]$ $[k^{\urcorner}t^h]$

Figure 8. For all stop-stop clusters, the amplitude of the signal was reduced to 0dB from after closure of the first stop to before release of the second.

case of perceptual metathesis. But if the subject decided it was a word and said *mixer*, that would be considered an example of metathesis.

The MEL program was used to run the experiment from a PC, collecting reaction time manual responses from a button box. The stimuli were played and the subjects' oral responses were recorded on a Sennheiser HMD410 headphone/microphone. A Quest Electronics Model 155 impulse precision sound level meter measured the stimuli for the comfortable listening level at approximately 75dB SPL with A weighting and F response, ±5dB depending on vowels and consonants. The near-speech reception threshold[9] was at approximately 40dB SPL, same conditions. Oral responses were recorded onto professional audio-tape at half-speed in order to fit an average 45min session on one side of a 60min tape.

Subjects were informed that the first task was to decide whether an English word[10] was spoken or not. They were to press the "YES" button with their right index finger if they thought the token was a word, otherwise they were to press the "NO" button with their left index finger. After making the lexical decision, the subjects were instructed to perform the second task of repeating aloud what they heard, as best they could.

The listeners performed the tasks individually in a sound-attenuated room. After a practice trial using a representative selection of word and non-word foils to ensure the subjects' comprehension of the task, the stimuli were randomly presented in six blocks. Subjects were allowed to pause after each block, and were given a rest break after the third block.

RT analyses of the CLL group were performed on the correct rejection of non-word targets (the metathesized words) as words. To determine what types of perceptual errors listeners made, their audio-tapes were transcribed auditorily and through the examination of spectrograms. Error analyses of the SRT group were performed on the metathesis and non-metathesis errors.

5.3. Results and discussion

RT analysis of CLL correct rejection of non-word targets Overall, CLL subjects were *slower* to identify non-word targets with optimal clusters ([tæski] from *taxi*) than those with non-optimal clusters ([wɪksi] from *whiskey*), as shown in Figure 9. There was a significant effect of optimality on RT (optimal cluster words were slower than non-optimal ones by 79ms, $F = 34.624$, $p < .05$), and obstruent type (stop-stop cluster words were slower than those with fricatives by 43ms, $F = 24.053$, $p < .05$). There was also an interaction between optimality and obstruent type ($F = 7.121$, $p < .05$). Further, there was a significant effect of obstruent ordering on RT (stop-fricatives are faster than fricative-stops by 59ms, and are faster than stop-stops by 72ms, $F = 16.765$, $p < .05$).

---

[9] 40dB SPL is the level that corresponded to SRT for most participants in Winters (2001).

[10] Subjects were instructed to treat compound words like *greenhouse*, *one-way*, and *ice cream* as single words—anything they would expect to find listed in the dictionary.

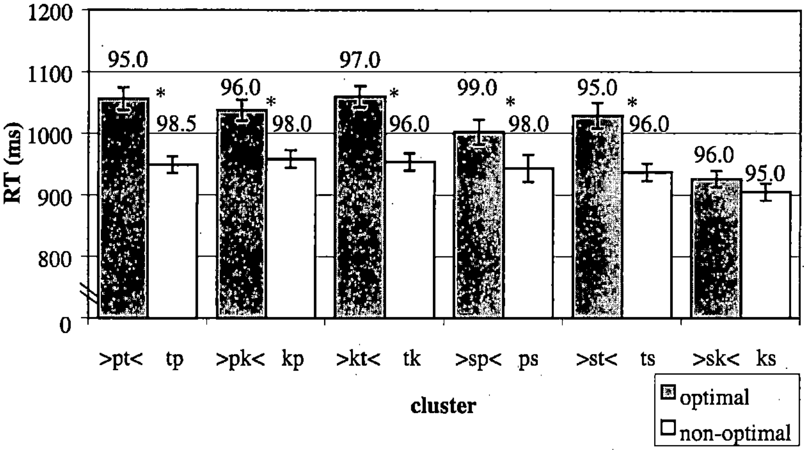


Figure 9. Mean reaction times, standard errors, and percentages correct for the correct rejections of targets in the auditory lexical decision task. Optimal clusters that were significantly slower than their non-optimal pairs are asterisked.

Although the prediction was that the non-optimal clusters would be more confusing, taking a longer time to respond "NO" to, the opposite result was generally found: Optimal clusters have a longer reaction time. This is in keeping with findings by Vitevitch and Luce (1999) in which listeners take longer to reject non-words that are word-like (having a high probability/density of segment sequences, i.e. they are phonetically similar to many words) than non-words that are not word-like. Since the optimal clusters occur in more English words than non-optimal clusters do, they are more word-like, and thus are harder to discount as words.

There is a significant effect of optimality for each pair ($p < .05$) except for >[sk]< and [ks]. Recall that [ks] is the only non-optimal sequence with a higher frequency of occurrence than its optimal pair. This result then could be a cluster frequency effect. However, if that were the case, the prediction would be that [ks] would have a significantly higher RT than >[sk]<. The solution is that there are both optimal perceptual clustering and cluster frequency effects. Since all the other optimal clusters occurred at least as much and usually much more than their non-optimal pairs, their higher frequency of occurrence gave a boost to subjects' performance which was already high based on perceptibility. However, since >[sk]< occurs less frequently than [ks], optimal >[sk]< did not gain this frequency boost.

Error analysis of SRT optimal vs. non-optimal clusters of non-word targets[11]
Figure 10 shows the number of metathesis and non-metathesis errors for optimal and non-optimal clusters for listeners in the speech reception threshold condition. Subjects

[11] See Appendix D for CLL group errors, and Appendix E for SRT group errors.

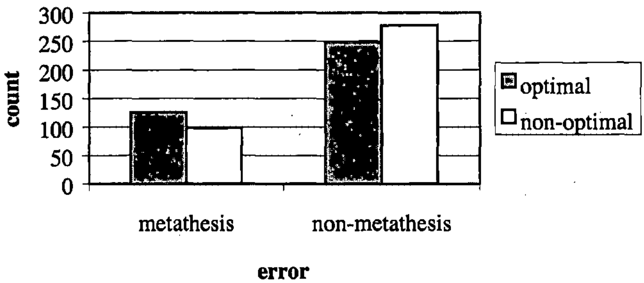


Figure 10. Types of errors for SRT optimal and non-optimal cluster target items. Optimal clusters had significantly more metathesis errors but fewer non-metathesis errors than non-optimal clusters.

made *more* metathesis errors on the non-word targets with optimal clusters (metathesizing them back into English words) than those with non-optimal clusters. Subjects made *fewer* non-metathesis errors on the optimal clusters than the non-optimal ones. This pattern is significant ($p < .05$). The interpretation of why the optimal clusters metathesized is as follows:

This experiment was attempting to cause listeners to metathesize *to* real words, not *from* real words, as it is attested in language. The results show, in effect, a "metathesis in reverse"—hearing good cues leads the listener back to the underlying form, instead of the underlying form metathesizing to result in good cues that will be preserved. Since the optimal clusters have better cues than non-optimal clusters, there is a higher probability that the listeners heard both obstruents in the optimal clusters correctly. For metathesis to occur, there need to be two obstruents to switch. Because the non-optimal clusters have poorer cues, it is likely that one or both consonants were not heard correctly, and therefore cannot be metathesized—other errors are made instead. If the subjects heard the optimal clusters, and heard enough of the rest of the word to narrow down the word choices, then a temporal change would result in a lexical item. Connine et al (1993) found that changing a few features of a phone can still lead to priming of the base word, so switching features could have similar effects. Since the majority of the real word sources of the targets had zero or one neighbors, if any word was activated during recognition it was more than likely to be one of those.

Error analysis of SRT manner features of non-word targets The number of metathesis and non-metathesis errors for stop-stop and fricative%stop (i.e. fricative-stop and stop-fricative) clusters are shown in Figure 11. Clusters with fricatives and stops were significantly less likely to metathesize than those composed solely of stops ($p < .05$). This supports Steriade's claim that the linear order of adjacent consonants that share manner features is highly non-salient.

Stop-stop clusters and fricative%stop clusters had the same amount of non-metathesis errors, which indicates that fricative%stop clusters are no less salient than

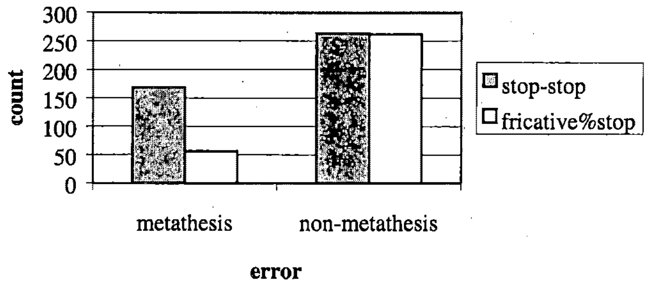


Figure 11. Types of errors for SRT stop-stop and fricative%stop cluster target items. Stop-stop clusters had significantly more metathesis errors than clusters with fricatives, but had roughly the same amount of non-metathesis errors.

stop-stop clusters. However, fricative%stop clusters caused fewer perceptual metathesis errors, indicating that their temporal ordering is more salient than that of stop-stop clusters. Judging from the fact that stops share more manner features than fricatives and stops, the more manner features two consonants share, the fewer cues there are to determine their order.

6. SUMMARY AND CONCLUSION

This study examined acoustic and perceptual cues in obstruent clusters in order to test the hypothesis that metathesis can be a process that maintains identification of the consonants involved. Clusters with poor cues may be susceptible to sound change, but if an obstruent with poor cues can switch to a position that improves its perceptibility, this optimal cluster has a better chance of preservation, as proposed in Hume 1998, 2001 and Steriade 2001. In English, most of the predicted optimal clusters were found to be more prevalent in the lexicon than non-optimal clusters. This could be proof that optimal clusters are more likely to be maintained.

In an auditory lexical decision task, there were effects of both optimality of cues, and frequency of clusters in the lexicon. For the clear listening level group, there was a slow rejection of targets with clusters that occur with high frequency in the lexicon. This usually was in tandem with the slow rejection of optimal clusters, except for >[sk]< and [ks], in which the non-optimal [ks] had a higher lexical frequency. For the speech reception threshold group, targets with optimal clusters were more likely to be perceptually metathesized and realized as the underlying words than targets with non-optimal clusters were because subjects are more likely to hear both consonants in optimal clusters. Clusters with fricatives and stops were less likely to be perceptually metathesized than clusters containing only stops, since the continuity of manner features in a cluster hinders perception of consonant order. Thus good cues indicating the transition between the obstruents in a cluster are important as well as cues into and out of the cluster.

In conclusion, the results of this study suggest that examining the lexicons of languages with metathesis in conjunction with following perceptual principles may provide explanations to some of the patterns observed in language sound systems. Although some of the perceptual findings will need to be adapted for the acoustics of a specific language (such as for languages that do not lenite /t/ postvocalically as in American English), in general, most good perceptual cues are language universal.

REFERENCES

Baayen, R.H., Piepenbrock, R., and Gulikers, L. (1995). *The CELEX Lexical Database* (Release 2) [CD-ROM]. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania [Distributor].

Bailey, C-J. N. (1970). Toward specifying constraints on phonological metathesis. *Linguistic Inquiry*, 1, 347-349.

Blumstein, S. E., and Stevens, K. N. (1979). Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants. *The Journal of the Acoustical Society of America*, 66, 1001-1017.

Bond, Z. S. (1971). Units in speech perception. *Working Papers in Linguistics*, 9, viii-112, Ohio State U.

Bregman, A. S. (1990). *Auditory Scene Analysis.* Cambridge, MA: MIT Press.

Connine, C. M., Blasko, D. G., and Titone, D. (1993). Do the beginnings of spoken words have a special status is auditory word recognition? *Journal of Memory and Language*, 32, 193-210.

Delattre, P. C., Liberman, A. M., and Cooper, F. S. (1955). Acoustic loci and transitional cues for consonants. *The Journal of the Acoustical Society of America*, 27, 769-773.

Fay, W. H. (1966). Temporal sequence in the perception of speech. *Janua Linguarium*, Series Minor 45.

Frisch, S. A., Large, N. R., and Pisoni, D. B. (2000). Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language*, 42, 481-496.

Goldinger, S. D. (1996). Auditory lexical decision. *Language and Cognitive Processes*, 11, 559-567.

Hume, E. (1998). The role of perceptibility in consonant/consonant metathesis. *Proceedings of the West Coast Conference on Formal Linguistics*, 17, 293-307.

Hume, E. (2001). Metathesis: Data, Motivation and Phonological Theory. In E. Hume, N. Smith and J. van de Weijer (eds.), *Surface Syllable Structure and Segment Sequencing*. Leiden, NL: HIL.

Hume, E., Johnson, K., Seo, M., and Tserdanelis, G. (1999). A cross-linguistic study of stop place perception. *Proceedings of the XIVth International Congress of Phonetic Sciences*, 2069-2072.

Hock, H. H. (1985). Regular metathesis. *Linguistics*, 23, 529-546.

Kewley-Port, D. (1983). Time-varying features as correlates of place of articulation in stop consonants. *The Journal of the Acoustical Society of America*, 73, 322-335.

Lahiri, A., Gewirth, L., and Blumstein, S. E. (1984). A reconsideration of acoustic invariance for place of articulation in diffuse stop consonants: Evidence from a cross-language study. *The Journal of the Acoustical Society of America*, 76, 391-404.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (ed.), *Cerebral Mechanisms in Behavior: The Hixon Symposium*, 112-146, New York: Wiley.

Luce, P. A. (1986). Neighborhoods of words in the mental lexicon. *Research on Speech Perception* Technical Report No. 6, Indiana U.

Luce, P. A., and Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 19, 1-36.

Miller, G. A., and Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *The Journal of the Acoustical Society of America*, 27, 338-352.

Nakao, T. (1986). Metathesis. In D. Kastovsky and A. Szwedek (eds.), *Linguistics across Historical and Geographical Boundaries: In Honour of Jacek Fisiak on the Occasion of His Fiftieth Birthday*, 547-556, New York: Mouton de Gruyter.

Öhman, S. E. G. (1965). Coarticulation in VCV utterances: Spectrographic Measurements. *The Journal of the Acoustical Society of America*, 39 1966 p.151-168.

Pickett, J. M. (1958). Perception of compound consonants. *Language and Speech*, 1(4), 288-304.

Pitt, M. A., and McQueen, J. M. (1998). Is compensation mediated by the lexicon? *Journal of Memory and Language*, 39, 347-370.

Pollack, I., Rubenstein, H., and Decker, L. (1959). Intelligibility of known and unknown message sets. *The Journal of the Acoustical Society of America*, 31, 273-279.

Remez, R. E. (2001). The multimodal nature of speech. In E. Hume and K. Johnson (eds.), *The Role of Speech Perception in Phonology*. New York: Academic Press (in press).

Savin, H. B. (1963). Word-frequency effect and errors in the perception of speech. *The Journal of the Acoustical Society of America*, 35, 200-206.

Seo, M., and Hume, E. (2001). A comparative OT account of metathesis in Faroese and Lithuanian. In E. Hume, N. Smith and J. van de Weijer (eds.), *Surface Syllable Structure and Segment Sequencing*. Leiden, NL: HIL.

Silva, C. M. (1973). Metathesis of obstruent clusters. *Working Papers in Linguistics*, 14, 77-84, Ohio State U.

Steriade, D. (2001). Directional asymmetries in place assimilation: A perceptual account. In E. Hume and K. Johnson (eds.), *The Role of Speech Perception in Phonology*. New York: Academic Press (in press).

Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics*, 17, 3-45.

Ultan, R. (1978). A typological view of metathesis. In J. H. Greenberg (ed.), *Universals of Human Language*, vol 2., 367-402. Stanford: Stanford UP.

Vitevitch, M. S., and Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, 40, 374-408.

Wang, M. D., and Bilger, R. C. (1973). Consonant confusion in noise: A study of perceptual features. *The Journal of the Acoustical Society of America*, 54, 1248-1266.

Winters, S. (2001). Putting place in its place: Evaluating place perception in VCCV sequences. In E. Hume, N. Smith and J. van de Weijer (eds.), *Surface Syllable Structure and Segment Sequencing*. Leiden, NL: HIL.

Wright, R. (1996). *Consonant Clusters and Cue Preservation in Tsou*. PhD dissertation. UCLA.

Wright, R. (2001). Perceptual cues in contrast maintenance. In E. Hume and K. Johnson (eds.), *The Role of Speech Perception in Phonology*. New York: Academic Press (in press).

Yost, W. A. (1994). *Fundamentals of Hearing: An Introduction* (3rd Ed.). New York: AP.

Appendix A. Targets and their source words.

Target	Word	Written and spoken word frequency	Spoken word frequency	Number of neighbors
>pt<				
nɪptɪk	nitpick	0	0	0
hæptɪn	hatpin	8	0	0
fʊptæθ	footpath	50	3	1
aʊptʊt	output	517	54	1
hiptʌmp	heat pump	0	0	0
ʃɑptʊt	shot-put	0	0	0
fʊptæd	footpad	3	0	0
hɑptɑt	hotpot	4	0	1
swipti	sweet pea	0	0	0
splɪpti	split pea	0	0	0
	Sum	582	57	3
>pk<				
stɑpkaɪl	stockpile	56	0	1
d͡ʒæpkɑt	jackpot	9	0	1
t͡ʃɪpki	chickpea	6	0	0
kɑpkɪt	cockpit	59	3	0
kræpkɑt	crackpot	10	0	1
stɑpkɑt	stockpot	3	0	0
t͡ʃɛpkɔɪnt	checkpoint	44	0	0
straɪpke	strike-pay	0	0	0
sɪpke	sick-pay	12	0	1
t͡ʃɔpkɪt	chalkpit	0	0	0
	Sum	199	3	4
>kt<				
aʊktʌm	outcome	379	16	0
waɪktæp	whitecap	6	0	2
suktes	suitcase	334	3	0
naɪktæp	nightcap	22	0	1
kæktɔl	catcall	6	0	1
aʊktæst	outcast	39	0	4
fruktek	fruitcake	8	0	0
frektɑr	freight car	0	0	0
striktɑr	streetcar	15	0	0
nʌktes	nutcase	0	0	1
	Sum	809	19	9

Target	Word	Written and spoken word frequency	Spoken word frequency	Number of neighbors
tp				
krɪtpɪk	cryptic	44	0	3
ʌtpaʊn	uptown	7	0	2
kætpɪv	captive	127	4	3
tɪtpo	tiptoe	104	0	2
pɛtpɔk	pep talk	0	0	1
ʌtpaɪt	uptight	26	3	3
bætpɪst	Baptist	100	1	1
rɛtpaɪl	reptile	110	0	0
ʌtpɚnd	upturned	48	0	1
sɛtpɚ	sceptre	25	0	5
	Sum	566	8	16
kp				
nækpɪn	napkin	124	1	0
taɪkpæst	typecast	1	0	0
wɪkpɔrd	whipcord	6	0	1
ʌkpip	upkeep	30	0	0
kʌkpek	cupcake	0	0	0
tɑkpot	topcoat	9	0	0
slɪkpot͡ʃ	slip-coach	0	0	0
pɑkpɔrn	popcorn	14	0	0
rɪkpɔrd	ripcord	1	0	2
zɪkpod	zip code	0	0	0
	Sum	185	1	3
tk				
vɪtkɚ	victor	178	10	2
ɑtkɪv	octave	35	3	2
spɛtkɚ	spectre	49	0	1
pɛtkɪn	pectin	26	0	1
lætkɪk	lactic	33	0	1
tætkəl	tactile	30	1	0
kɑtkel	cocktail	179	2	0
lɛtkɚn	lectern	40	2	1
trætkɚ	tractor	191	7	1
lætkos	lactose	8	0	0
	Sum	769	25	9

Appendix A. cont.

>sp<				
d͡ʒɪspi	gipsy	96	2	1
ʌspɚd͡ʒ	upsurge	56	3	0
tɪspi	tipsy	18	0	1
ɛspəm	Epsom	8	0	0
tɑspɔɪl	topsoil	32	0	0
tɑspaɪd	topside	13	0	0
pɛspɪn	pepsin	0	0	0
stɛspʌn	stepson	6	0	0
næspæk	knapsack	17	0	0
sospʌdz	soapsuds	8	0	0
	Sum	254	5	2

ps				
d͡ʒæpsɚ	jasper	28	0	0
aɪpsɪk	icepick	1	0	2
krɪpsi	crispy	5	0	4
æpsən	aspen	15	0	0
tipsun	teaspoon	65	2	1
dɛpsət	despot	21	0	0
prɑpsɚ	prosper	106	5	2
hɑpsɪs	hospice	6	0	0
gripsent	greasepaint	8	0	0
sɛpsul	cesspool	6	1	0
	Sum	261	8	9

>st<				
pistə	pizza	34	1	0
fʊsti	footsie	1	0	1
fʊstɔr	footsore	4	0	1
aʊstɛt	outset	100	7	2
kɚsti	curtsy	44	0	3
kæstəp	catsup	8	0	0
nɑsti	Nazi	372	9	1
skɪstɔɪd	schizoid	7	0	0
d͡ʒɛstəm	jetsam	10	0	0
stestaɪd	stateside	9	0	1
	Sum	589	17	9

ts				
pɑtsə	pasta	36	2	2
itsɚ	Easter	265	19	5
gʌtso	gusto	32	0	2
tetsi	tasty	56	1	5
hetsæk	haystack	21	1	0
kɔtsɪk	caustic	27	0	2
mʌtsɚ	muster	104	0	14
spætsɪk	spastic	11	0	1
d͡ʒɛtsɪŋ	jesting	2	0	1
plætsɚd	plastered	6	0	2
	Sum	560	23	34

>sk<				
wæski	waxy	24	0	7
tæski	taxi	645	27	7
hæskɔ	hacksaw	5	0	0
pɪski	pixie	1	0	4
dɪski	Dixie	6	0	4
mɪskɚ	mixer	29	1	2
ɛskɪt	exit	253	23	2
tɑskɪk	toxic	106	0	2
flæskən	flaxen	2	0	0
ɛskaɪz	excise	16	1	3
	Sum	1087	52	31

ks				
wɪksi	whiskey	623	11	4
haʊksot	housecoat	21	0	1
hʌksi	husky	45	0	6
pɛksi	pesky	1	0	0
dɪkso	disco	150	2	1
mʌksi	musky	8	0	8
ɛksɔrt	escort	138	1	1
kæksɪt	casket	39	0	1
frɪksi	frisky	11	1	3
vɪksəs	viscous	23	10	1
	Sum	1059	25	26

Appendix B. English word foils.

>pt<	**tp**	**>sp<**	**ps**
riptide	bit part	peace pipe	flip side
styptic	foot-pound	lisping	topsail
striptease	gatepost	crosspiece	dropsy
optic	outpost	waspish	gypsum
sceptic	jetpack	space probe	typeset
uptake	footprint	tailspin	keepsake
raptly	lightproof	misprint	ripsaw
claptrap	hotplate	spoilsport	lapse rate
aptly	waste-pipe	sunspot	upswing
sculptor	dustpan	homespun	campsite
>pk<	**kp**	**>st<**	**ts**
upcast	stickpin	blast-off	jet set
stopcock	neckpiece	mastiff	hot seat
tipcart	backpack	coster	ritzy
slipcase	bookplate	nesting	wet suit
dropkick	spark-plug	taster	pretzel
shopkeep	pickproof	all-star	heartsick
pipeclay	leakproof	shoestring	shirtsleeve
bumpkin	shockproof	brainstorm	pint-sized
pumpkin	inkpad	tombstone	Scotsman
trumpcard	inkpot	limestone	statesman
>kt<	**tk**	**>sk<**	**ks**
backtalk	flatcar	Peace Corps	rock-salt
folktale	oatcake	play-school	hoaxer
ductile	gatecrash	mascot	quicksand
proctor	yacht-club	whiskers	axle
shock troops	shortcake	basket	accent
backtrack	nightclub	icecube	laxly
actress	outcry	bearskin	waxwork
spectral	shift key	dunce cap	locksmith
arctic	test case	briskness	blacksmith
tactful	postcard	task-force	Oxsford

Appendix C. Non-word foils and their source words with no neighbors and zero frequency.

>pt<	
ɛpta^{u}nd	egg-bound
brɔpted͡ʒ	broad gauge
smoptɑm	smoke-bomb
nɛptænd	neckband
lʌptel	lugsail
bɚptɔg	bird dog
stra^{ɪ}pta^{u}nd	strikebound
sɪptɚθ	sick-berth
sɪptɔl	sick call
bɚptid	birdseed

>pk<	
flʌpka^{ɪ}d	flood-tide
rʊpkir	root beer
wepkænd	wave band
tʌpkɑp	tuck-shop
bɛpkost	bedpost
drɪpka^{ɪ}s	drift-ice
drɪpkɪd͡ʒ	driftage
pɑpka^{u}nd	potbound
ropkʊk	road-book
flæpkɑp	flattop

>kt<	
rɪkted͡ʒ	rib cage
wʊstʌlp	wood-pulp
roktɛns	road-sense
diktol	deedpoll
blʌktæŋk	blood bank
ʃɑktɚl	shop-girl
fespe	fete-day
pɑpka^{u}nd	potbound
bɛktɔr	bedsore
hɑktol	hop-pole

tp	
slutpælv	sluice-valve
trʌtpɑrm	truck farm
stɑtpɑr	stockcar
blætpæp	black cap
prɛtpɑks	press-box
pɪtpot	pigboat
fɑtpæŋk	fogbank
brɪtpɪln	brickkiln
gritpʌn	grease-gun
lutpɑks	loosebox

kp	
glokpɪʃ	globefish
bra^{ɪ}kpek	bridecake
frukpæt	fruit bat
brɛkpɔrd	breadboard
slɑkpol	slop bowl
klɔkpæns	clog-dance
ʃɑkpɛl	shop-bell
stʌkpʊk	stud-book
ʃɑkpɔ^{ɪ}	shop-boy
spikpɑp	speed-cop

tk	
tetkɛk	tape deck
a^{ɪ}tkit	ice sheet
a^{ɪ}tkot	iceboat
ha^{u}tkɔg	housedog
t͡ʃɔtkɑr	choc-bar
pletkɛt	place-bet
fɑtka^{u}nd	fogbound
rɑtkʌn	rock bun
sætkʌt	sackbut
pɑtkɑp	popshop

Appendix C. cont.

>sp<		**ps**	
blʌspa^{ʊ}nt	blood count	kwapsɑ	kwacha
kæspɛsk	cash desk	t͡ʃopsæmp	choke-damp
dɛspæks	death tax	ɛpsek	egg-shake
wa^{ɪ}spa^{ɪ}	wise guy	brɪpsild	brickfield
flæspʌn	flashgun	spopsev	spokeshave
pʌspo	Pashto	siskend͡ʒ	sea change
sispa^{ɪ}m	seedtime	bɚptɔg	bird dog
na^{ɪ}stɛl	night-bell	stɑpsɪʃ	stockfish
fespe	fete-day	dɪpsɔŋ	diphthong
sæspɔrd	sash-cord	hipsild	heat shield
>st<		**ts**	
bla^{ɪ}stəm	blithesome	pa^{ɪ}tsʊl	pipeful
ta^{ɪ}stɑrn	tithe-barn	a^{ɪ}tsɔl	icefall
pɑstɔ^{ɪ}	pot-boy	stɪtsɪft	stick shift
kla^{ʊ}stæŋk	cloudbank	a^{ɪ}tsild	icefield
skwɑstɑr	squad car	a^{ɪ}tso	ice-show
sa^{ʊ}stɔ	southpaw	witsɔrm	weak form
listʌd	leaf-bud	pʌtsɔg	pug-dog
wʊstʌlp	wood-pulp	dɔtsez	dog-days
trestæp	trade gap	hɑtsild	hop-field
na^{ɪ}stɛl	night-bell	rætse	rag-day
>sk<		**ks**	
fla^{ɪ}skɛk	flight deck	floksɑrt	flowchart
pruskit	proof sheet	pʊksʊl	pushful
fɪskest	fishpaste	slʌksʌnd	slush fund
skɪskæn	skidpan	ska^{ɪ}ksæk	skyjack
striskɚl	street-girl	slɑksɑp	slop-shop
siskend͡ʒ	sea change	tʌksʊl	tubful
na^{ɪ}skɪft	night shift	sa^{ɪ}kses	side-face
mʌskæk	mudpack	ha^{ɪ}kser	high chair
piskɪk	peachick	ha^{ɪ}ksʌmp	high jump
pæskɔl	pat-ball	ʃiksɪp	sheepdip

Appendix D. Spoken errors of CLL group.

Error	Count	% of errors	% of targets
metathesis	24	10.71	1.00
anticipatory assimilation	107	47.77	4.46
perseveratory assimilation	7	3.13	0.29
delete C1	23	10.27	0.96
delete C2	0	0.00	0.00
change feature of C1	51	22.77	2.13
change feature of C2	1	0.45	0.04
insertion[12]	9	4.02	0.38
other	2	0.89	0.08
TOTAL	224		9.33

2400 total targets
20 subjects

CLUSTER	metath	antic	persev	del 1	del 2	ch 1	ch 2	insert	other	TOTAL
>pt<	9	31	1	0	0	6	0	0	0	47
tp	0	8	0	2	0	4	0	0	0	14
>pk<	3	0	1	0	0	3	0	0	0	7
kp	4	12	1	0	0	8	0	1	1	27
>kt<	6	33	0	1	0	5	0	0	1	46
tk	1	23	3	20	0	23	0	0	0	70
>sp<	0	0	0	0	0	0	1	0	0	1
ps	1	0	0	0	0	0	0	3	0	4
>st<	0	0	1	0	0	0	0	3	0	4
ts	0	0	0	0	0	2	0	0	0	2
>sk<	0	0	0	0	0	0	0	1	0	1
ks	0	0	0	0	0	0	0	1	0	1
>TT<	18	64	2	1	0	14	0	0	1	100
TT	5	43	4	22	0	35	0	1	1	111
>sT<	0	0	1	0	0	0	1	4	0	6
Ts	1	0	0	0	0	2	0	4	0	7
optimal	18	64	3	1	0	14	1	4	1	106
non-optimal	6	43	4	22	0	37	0	5	1	118

[12] Errors classified as insertions may have other errors besides the insertion.

Appendix E. Spoken errors of SRT group.

Error	Count	% of errors	% of targets	
metathesis	225	29.96	18.75	
anticipatory assimilation	45	5.99	3.75	
perseveratory assimilation	10	1.33	0.83	
delete C1	19	2.53	1.58	
delete C2	22	2.93	1.83	
change feature of C1	101	13.45	8.42	
change feature of C2	113	15.05	9.42	
insertion	121	16.11	10.08	
other	95	12.65	7.92	
TOTAL	751		62.58	1200 total targets 10 subjects

CLUSTER	metath	antic	persev	del 1	del 2	ch 1	ch 2	insert	other	TOTAL
>pt<	28	1	2	2	1	4	16	11	10	75
tp	25	11	1	2	1	7	22	1	11	81
>pk<	10	5	2	2	0	10	3	8	2	42
kp	44	6	2	1	1	7	8	8	4	81
>kt<	50	5	1	0	1	9	0	5	11	82
tk	12	14	0	2	1	24	5	9	5	72
>sp<	21	2	1	1	3	2	24	6	15	75
ps	6	1	0	0	1	9	9	15	11	52
>st<	14	0	0	2	8	8	3	14	6	55
ts	9	0	1	3	3	15	12	15	11	69
>sk<	3	0	0	2	1	3	10	20	6	45
ks	3	0	0	2	1	3	1	9	3	22
>TT<	88	11	5	4	2	23	19	24	23	199
TT	81	31	3	5	3	38	35	18	20	234
>sT<	38	2	1	5	12	13	37	40	27	175
Ts	18	1	1	5	5	27	22	39	25	143
optimal	126	13	6	9	14	36	56	64	50	374
non-optimal	99	32	4	10	8	65	57	57	45	377